

## FIGURE 1

CGGACGCGTGGGACCCATACTTGCTGGTCTGATCCATGCACAAGGCGGGGCTGCTAGGCCTC  
TGTGCCCGGGCTTGGAATTCGGTGC GGATGGCCAGCTCCGGGATGACCCGCCGGGACCCGCT  
CGCAAATAAGGTGGCCCTGGTAACGGCCTCCACCAGCGGGATCGGCTTCGCCATCGCCCGGC  
GTTTGGCC CAGGACGGGGCCCATGTGGTCGT CAGCAGCCGGAAGCAGCAGAATGTGGACCAG  
GCGGTGGCCACGCTGCAGGGGGAGGGGCTGAGCGTGACGGGCACCGTGTGCCATGTGGGGAA  
GGCGGAGGACCGGGAGCGGCTGGTGGCCACGGCTGTGAAGCTTCATGGAGGTATCGATATCC  
TAGTCTCCAATGCTGCTGTCAACCCCTTTCTTTGGAAGCATAATGGATGTCACTGAGGAGGTG  
TGGGACAAGACTCTGGACATTAATGTGAAGGCCCCAGCCCTGATGACAAAGGCAGTGGTGCC  
AGAAATGGAGAAACGAGGAGGCGGCTCAGTGGTGATCGTGTCTTCCATAGCAGCCTTCAGTC  
CATCTCCTGGCTTCAGTCCTTACAATGT CAGTAAACAGCCTTGCTGGGCCTGACCAAGACC  
CTGGCCATAGAGCTGGCCCCAAGGAACATTAGGGTGAAGTGCCTAGCACCTGGACTTATCAA  
GACTAGCTTCAGCAGGATGCTCTGGATGGACAAGGAAAAAGAGGAAGCATGAAAGAAACCC  
TGCGGATAAGAAGGTTAGGCGAGCCAGAGGATTGTGCTGGCATCGTGTCTTCTGTGCTCT  
GAAGATGCCAGCTACATCACTGGGGAAACAGTGGTGGTGGGTGGAGGAACCCCGTCCCGCCT  
CTGAGGACCGGGAGACAGCCACAGGCCAGAGTTGGGCTCTAGCTCCTGGTGCTGTTCTCTGC  
ATTCACCCACTGGCCTTTCCACCTCTGCTCACCTTACTGTTACCTCATCAAATCAGTTCT  
GCCCTGTGAAAAGATCCAGCCTTCCCTGCCGTCAAGGTGGCGCTCTTACTCGGGATTCTGTCT  
GTTGTTGTGGCCTTGGGTAAAGGCCTCCCCTGAGAACACAGGACAGGCCTGCTGACAAGGCT  
GAGTCTACCTTGGCAAAGACCAAGATATTTTTCTGGGCCACTGGTGAATCTGAGGGGTGA  
TGGGAGAGAAGGAACCTGGAGTGAAGGAGCAGAGTTGCAAATTAACAGCTTGCAAATGAGG  
TGCAAATAAAATGCAGATGATGCGCGGCTTTGAAAAAAAAA

# 2023

```
><subunit 1 of 1, 278 aa, 1 stop
```

MHKAGLLGLCARAWNSVRMASSGMRTRDPLANKVALVTASTDGIGFAIARRLAQDGAHVVVV  
SRKQQNVQDQAVATLQGEGLSVTGTVCHVKGAEADRERLVATAVKLHGGIDILVSNAAVNPFFG  
SIMDVTEEVDKTLTDINVKAPALMTKAVVPEMEKRGGSVVIVSSIAAFSPSPGFSPPYNVSK  
TALLGLTKTLAIELAPRNI RVNCLAPGLIKTFSRMLWMDKEKEESMKETLIRRLGEPEDC  
AGIVSFLCSEDASYITGETVVVGGGTPSRL

Signal peptide:

N-glycosylation site.

amino acids 183-186

### N-myristoylation sites.

amino acids 43-48, 80-85, 191-196, 213-218, 272-277

Microbodies C-terminal targeting signal.

amino acids 276-278



## FIGURE 4

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA47465  
<subunit 1 of 1, 830 aa, 1 stop  
<MW: 95029, pI: 8.26, NX(S/T): 2  
MEQYKLSQSDRLREQQEEMVELRLRLVLRPGWGLRLNGLPPGSFVPRHTAPLGGAHAHV  
LGMVFPACLPFDEVSGEQRGEQVTNGREAGAEELLTEVNRLLGSGSSAASEEEEEEEPPRRTL  
HLRRNRISNCSQRAGARPGSLPERKGPCLCEELDAAIPGSAVGGSKARVQARQVPPATAS  
EWRLAQAAQQKIRELAINIRMKEELIGELVRTGKAAQALNRQHSQRIRELEQEAQVRAELSE  
GQRQLRELEGKELQDAGERSRLQEFRRRVAAAQSQVQVLKEKKQATERLVLSAQSEKRLQE  
LERNVQLMRQQQGQLQRLRLREETEQRRLLEAEMSKRQHRVKELELKHEQQQKILKIKTEEIA  
AFQRRRSNGSGSVSLEQQQKIEEQKKWLDQEMEKVLQRRALEELGEELHKKREAILAKKE  
ALMQEKTGLESKRRLSSQALNEDIVRVSSRLEHLEKELSEKSGQLRQGSAAQSQQQIRGEIDS  
LRQEKDSSLKQRLRIDGKLRQSSLLSPEEERTLFQLDEAIEALDAAIEYKNEAITCRQVRLR  
ASASLLSQCEMNLMAKLSYLSSETRALLCKYFDKVVTLREEQHQQQIAFSELEMQLEEQQR  
LVYVLEVALERQRLQLEMDRQLTLQQKEHEQNMQLLLQQSRDHLGEGGLADSRQYFARIQALEK  
ELGRYMWINQELKQKLGGVNAVGHSGGEEKRSLCSEGRQAPGNEDELHLAPELLWLSPLTEG  
APRTREETRDVLVHAPLPLTWKRSSLCGEEQGSPEELRQREAAEPLVGRVLPVGEAGLPWNFG  
PLSKPRRELRRASPGMIDVRKNPL

### Important features:

#### Leucine zipper pattern.

amino acids 557-579, 794-815

#### N-glycosylation sites.

amino acids 133-136, 383-386

#### Kinesin related protein Kif-4 Coiled-coil domain:

amino acids 231-672

0966034.052501

## FIGURE 5

ATTCTCCTAGAGCATCTTTGGAAGCATGAGGCCACGATGCTGCATCTTGGCTCTTGTCTGCT  
GGATAACAGTCTTCCTCCTCCAGTGTTCAAAAGGAACCTACAGACGCTCCTGTTGGCTCAGGA  
CTGTGGCTGTGCCAGCCGACACCCAGGTGTGGGAACAAGATCTACAACCCCTTCAGAGCAGTG  
CTGTTATGATGATGCCATCTTATCCTTAAAGGAGACCCGCCGCTGTGGCTCCACCTGCACCT  
TCTGGCCCTGCTTTGAGCTCTGCTGTCCCGAGTCTTTTGGCCCCCAGCAGAAGTTTCTTTGTG  
AAGTTGAGGGTTCTGGGTATGAAGTCTCAGTGTCACTTATCTCCCATCTCCCGAGCTGTAC  
CAGGAACAGGAGGCACGTCTGTACCCATTAAAAACCCAGGCTCCACTGGCAGACGGCAGAC  
AAGGGGAGAAGAGACGAAGCAGCTGGACATCGGAGACTACAGTTGAACCTTCGGAGAGAAGCA  
ACTTGACTTCAGAGGGATGGCTCAATGACATAGCTTTGGAGAGGAGCCCAGCTGGGGATGGC  
CAGACTTCAGGGGAAGAATGCCTTCCTGCTTCATCCCCTTTCCAGCTCCCCTTCCCCTGAG  
AGCCACTTTTCATCGGCAATAAAATCCCCACATTTACCATCT

33866034.052501

## **FIGURE 6**

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA57700
><subunit 1 of 1, 125 aa, 1 stop
><MW: 14198, pI: 9.01, NX(S/T): 1
MRPRCCILALVCWITVFLQCSKGTTDAPVGSLWLCQPTPRCGNKIYNPSEQCCYDDAILS
LKETRRCGSTCTFWPCFELCCPESFGPQQKFLVKLRVLGMKSQCHLSPISRSCTRNRRHVLYP
```

### **Important features:**

#### **Signal sequence**

amino acids 1-21

#### **N-myristoylation sites.**

amino acids 33-39, 70-76

09866034-052501

## FIGURE 7

CCCACGCGTCCGCCCCACGCGTCCGGGTGCCACTCGCGCGCCGGCGCGCTCCGGGCTTCTCT  
TTTCCCTCCGACGCGCCACGGCTGCCCAGACATTCCGGCTGCCGGGTCTGGAGAGCTCCCCG  
AACCCTCCGCGGAGAGGAGCGAGGCGGCGCCAGGGTGGCCCCCGGGGCGCGCTTGGTCTCG  
GAGAAGCGGGGACGAGGCCGAGGATGAGCGACTGAGGGCGACGCGGGGCACTGACGCGAGTT  
GGGGCCGCGACTACCGCGAGCTGACAGCGCGATGAGCGACTCCCCAGAGACGCCCTAGCCCG  
GTGTGCGCGCCAGGCGGAGCGCGCAGGTGGGGCTGGGTGTTAGTGGTCCGCCCCACGCGGG  
TCGCGCGCCGGCCAGGATGGGCGCTGGCAACCGGGCCCGCGCCCGCGCTGCTACCCCTG  
CGCCCGCTGCGAGCCCGGCGTCCGGCCCCGCGCCTGCGCTCATGAGCGGCGGCTCCCGGCTG  
GCGGCGGCGCGCCCCGGGCTGTGAATGCGACTCGCCCCCGGCGCGCTCCCGCCCGCCCC  
GCCCCCGGGACGTGGTAGGGGATGCCAGCTCCAATGCGATGGCAGTTGGCGCGCTCTCCA  
GTTCCCTCTGGTCACCTGCTGCCTGATGGTGGCTCTGTGAGTCCGAGCATCCCGTGGAG  
AAGCTGGCCAGGACCCAGAGCAGCGGGCCAGGAGAAGCGTGAGCACGCCACTCGGGACGG  
CCCGGGGCGGGTGAACGAGCTCGGGCGCCCGCGCAGGGACGAGGGCGGACGCGCCGGGACT  
GGAAGAGCAAGAGCGGCCGTGGGCTGCCGGCCGTGAGCCGTGGAGCAAGCTGAAGCAGGCC  
TGGGTCTCCAGGGCGGGGGCGCCAAGGCCGGGGATCTGCAAGTCCGGCCCCGCGGGGACAC  
CCCGCAGGCGGAAGCCTGCGCCGAGCCGCCAGGACGCGATTGGCCCGAACTCGCGCCCA  
CGCCCGAGCCACCCGAGGAGTACGTGTACCCGACTACCGTGGCAAGGGCTGCGTGGACGAG  
AGCGGCTTCGTGTACGCGATCGGGGAGAAGTTCGCGCCGGGCCCCCTCGGCCTGCCGTGCCT  
GTGCACCGAGGAGGGGCCGTGTGCGCGAGCCGAGTGCCCGAGGCTGCACCCGCGCTGCA  
TCCACGTCGACACGAGCCAGTGCTGCCCGAGTGCAAGGAGAGGAAGAATACTGCGAGTTC  
CGGGGCAAGACCTATCAGACTTTGGAGGAGTTCGTGGTGTCTCCATGCGAGAGGTGTCGCTG  
TGAAGCCAACGGTGAGGTGCTATGCACAGTGTGAGCGTGTCCCAGACGGAGTGTGTGGACC  
CTGTGTACGAGCCTGATCAGTGCTGCCATCTGCAAAAATGGTCCAACTGCTTTGCAGAA  
ACCGCGGTGATCCCTGCTGCGAGAGAAGTGAAGACTGACGAGTGACCATATGCCACTGTAC  
TTATGAGGAAGGCACATGAGAAATCGAGCGCAGGCCATGTGCACGAGACATGAATGCAGGC  
AAATGTAGACGCTTCCAGAAACACAACTCTGACTTTTTCTAGAACATTTTACTGATGTGAA  
CATTCTAGATGACTCTGGGAACTATCAGTCAAAGAAGACTTTTGATGAGGAATAATGGAAAA  
TTGTTGGTACTTTTCTTTTCTTGATAACAGTTACTACAACAGAAGGAAATGGATATATTT  
AAAACATCAACAAGAACTTTGGGCATAAAATCCTTCTCTAAATAAATGTGCTATTTTCAG  
TAAGTACACAAAAGTACACTATTATATATCAAATGTATTTCTATAATCCCTCCATTAGAG  
CTTATATAAGTGTTTTCTATAGATGCAGATTAAAAATGCTGTGTTGTCAACCGTCAAAAAA  
AAAAAAAAAAAAAAAAAAAA

CGGCGTCCGCCCCACGCGTCCGGGTGCCACTCGCGCGCCGGCGCGCTCCGGGCTTCTCT

## FIGURE 8

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68818
><subunit 1 of 1, 325 aa, 1 stop
><MW: 35296, pI: 5.37, NX(S/T): 0
MPSSTAMAVGALSSSLVTCCLMVALLCSPSIPLEKLAQAPQPGQEKREHATRDGPGRVNEL
GRPARDEGGSGRDWKS KSGRGLAGREPWSKLKQAWVSQGGGAKAGDLQVRPRGDTTPQAEALA
AAQDAIGPELAPTPEPPPEEYVYPDYRGKGCVDSESGFVYAIGKEFAPGPSACPCLCTEEGPL
CAQPECPLRHPRCIHVDTSQCCPQCKERKNYCEFRGKTYQTLEEFVVSPECRCRCEANGEVL
CTVSACPQTECVDPVYEFDQCCPICKNGPNCFAETAVIPAGREVKTDECTICHCTYEETWR
IERQAMCTRHECRQM
```

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-27

#### **Transmembrane domain:**

amino acids 11-30

#### **Glycosaminoglycan attachment site.**

amino acids 80-83

#### **N-myristoylation sites.**

amino acids 10-15, 102-107, 103-108

#### **Cell attachment sequence.**

amino acids 114-117

#### **EGF-like domain cysteine pattern signature.**

amino acids 176-187

0956034-052501



## FIGURE 9

CAGCCACAGACGGGTCA**ATG**AGCGCGGTATTACTGTGGCCCTCCTGGGGTTCATCCTCCAC  
TGCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCC  
GACCTACCCCGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGA  
CACGTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGTCTCTCAAGGGCTGCACGG  
AGGCCAAGGACCAGGAGCCCGCTCACTGAGCACCGGATGGGCCCCGGCCTCTCCCTGATC  
TCCTACACCTTCGTGTGCCGCCAGGAGGACTTCTGCAACAACCTCGTTAACTCCCTCCCGCT  
TTGGGCCCCACAGCCCCAGCAGACCCAGGATCCTTGAGGTGCCAGTGTGCTTGTCTATGG  
AAGGCTGTCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGAT  
GGCCTCCTCAGGCTCAGGGGAGGAGGCATCTTCTCCAATCTGAGAGTCCAGGGATGCATGCC  
CCAGCCAGGTTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCCGGGTATGACTGAGA  
ACTGCAATAGGAAGATTTTCTGACCTGTTCATCGGGGGACCACTTATGACACACGGAAC  
TTGGCTCAAGAACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGCAGGT  
GTGTGAGGAGACGCTGCTGCTCATAGATGTAGGACTCACATCAACCTTGGTGGGGACAAAAG  
GCTGCAGCACTGTTGGGGCTCAAAATCCCAGAAGACCACCATCCACTCAGCCCCCTCTGGG  
GTGCTTGTGGCCTCCTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAG  
CAGCGTTCTGTGTAACCTCCCTCCCTCAAGCTGCCCCGTGCCAGGAGACCGGCAGTGTCT  
CTACCTGTGTGCAGCCCTTGGAACTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGG  
GGCGCCACTCATTGTTATGATGGGTACATTTCATCTCTCAGGAGGTGGGCTGTCCACAAAAT  
GAGCATTACGGGCTGCGTGGCCCAACCTTCCAGCTTCTTGTGAACCACACACAGACAAATCG  
GGATCTTCTCTGCGCGTGAGAAGCGTGATGTGCAGCCTCCTGCCCTCTCAGCATGAGGGAGGT  
GGGGCTGAGGGCCTGGAGTCTCTCACTTGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTG  
GTGGGGAGTGGTTTGCCCTTCTCTGCT**TAA**CTCTATTACCCCCACGATTCTTACCCTGCTGA  
CCACCCACACTCAACCTCCCTCTGACCTCATAACCTAATGGCCTTGACACAGATTCTTTTC  
CCATCTGTCCATGAATCATCTTCCCCACACAATCATTCATATCTACTCACCTAACAGCA  
ACACTGGGGAGAGCCTGGAGCATCCGACTTGGCCTATGGGAGAGGGGACGCTGGAGGAGTG  
GCTGCATGTATCTGATAATACAGACCCTGTCCTTTCA

CCGCGCTTCTGCT

## **FIGURE 10**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59847
><subunit 1 of 1, 437 aa, 1 stop
><MW: 46363, pI: 6.22, NX(S/T): 3
MSAVLLLLALLGFILPLPGVQALLCQFGTVQHVWVSDLPQRQWTPKNTSCDSGLGCQDTLMLI
ESGPQVSLVLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP
PADPGSLRCPVCLSMEGCLEGTTEEICPKGTTHCYDGLLRRLRGGGIFSNLRVQGCMPPQGCN
LLNGTQEIGFVGMTENCNRKDPLTCHRGTTIMTHGNLAQEPTDWTTSNTEMCEVGQVCQETL
LLIDVGLTSTLVGTGKCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN
SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLGGGLSTKMSIQGC
VAQPSSFLLNHRQIGIFSAREKRDVQPPASQHEGGGAEGLESLTWGVGLALAPALWWGVVCPSC
```

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-15

#### **Transmembrane domain:**

amino acids 243-260

#### **N-glycosylation sites.**

amino acids 46-49, 189-192, 382-385

#### **Glycosaminoglycan attachment sites.**

amino acids 51-54, 359-362

#### **N-myristoylation sites.**

amino acids 54-59, 75-80, 141-146, 154-159, 168-173, 169-174,  
198-203, 254-259, 261-266, 269-274, 284-289, 333-338, 347-352,  
360-365, 361-366, 388-393, 408-413, 419-424

## FIGURE 11

CGACGATGCTACGCGCGCCCGGCTGCCTCCTCCGGACCTCCGTAGCGCCTGCCGCGGGCCCTG  
GCTGCGGCGCTGCTCTCGTCGCTTGCGCGCTGCTCTCTCTAGAGCCGAGGGACCCGGTGGC  
CTCGTCGCTCAGCCCCATTTTCGGCACCAAGACTCGCTACGAGGATGTCAACCCCGTGCTAT  
TGTCGGGCCCCGAGGCTCCGTGGCGGGACCTGAGCTGCTGGAGGGGACCTGCACCCCGGTG  
CAGCTGGTGCGCCCTCATTGCGCCACGGCACCCGCTACCCACGGTCAAACAGATCCGCAAGCT  
GAGGCAGCTGCACGGGTGCTGTCAGGCCGCGGGTCCAGGATGGCGGGGCTAGTAGTACCG  
GCAGCCGCGACCTGGGTGCAGCGCTGGCCGACTGGCCTTTGTGGTACGCGGACTGGATGGAC  
GGGCAGCTAGTAGAGAAGGACGGCAGGATATGCGACAGCTGGCGCTGCGTCTGGCCTCGCT  
CTTCCCGGCCCTTTTCAGCCGTGAGAACTACGGCCGCTGCGGTCTATCACAGGTTCCAAGC  
ACCGCTGCATGGATAGCAGCGCGCCCTTCTGTCAGGGGCTGTGGCAGCACTACCACCTTGGC  
TTGCCGCCGCGGACGTCGCAGATATGGAGTTTGGACCTCCAACAGTTAATGATAAACTAAT  
GAGATTTTTTGATCACTGTGAGAAGTTTTAACTGAAGTAGAAAAAATGCTACAGCTCTTT  
ATCAGCTGGAAGCCTTCAAACTGGACCAGAAATGCAGAACATTTAAAAAAGTTGCAGCT  
ACTTTGCAAGTGCCAGTAAATGATTAAATGCAGATTTAATTCAGTAGCCCTTTTTCACCTG  
TTCATTTGACCTGGCAATTAAAGGTGTTAAATCTCCTTGGTGTGATGTTTTTGACATAGATG  
ATGCAAGGTATTAGAAATTTAAATGATCTGAAACAATATTGGAAAAAGGATATGGGTAT  
ACTATTAACAGTCGATCCAGCTGCACCTTGTTTCAGGATATCTTTCAGCACTTGGACAAAGC  
AGTTGAACAGAAACAAAGGTCAGCCCAATTTCTTCCAGTCATCCTCCAGTTTGGTCATG  
CAGAGACTCTTCTTCCACTGCTTCTCTCATGGGCTACTCAAAGACAAGGAACCCCTAACA  
GCGTACAATTACAAAAACAAATGCATCGGAAGTTCGGAAGTGGTCTCATTGTACCTTATGC  
CTCGAACCTGATATTTGTGCTTTACCACTGTGAAAATGCTAAGACTCCTAAAGAACAAATCC  
GAGTGCAGATGTTATTAATGAAAAGGTGTACCTTTGGCTTACTCACAAGAACTGTTTCA  
TTTTATGAAGATCTGAAGAACCACTACAAGGACATCCTTCAGAGTTTCAGAACCAAGTGAAGA  
ATGTGAATTAGCAAGGGCTAACAGTACATCTGATGAACATATGAGTAACTGAAGAACATTTTT  
AATTTCTTTAGGAATCTGCAATGAGTGATTACATGCTTGTAAATAGGTAGGCAATTCCTTGATT  
ACAGGAAGCTTTTATATTACTTGAGTATTTCTGTCTTTTCACAGAAAAACATTGGGTTTCTC  
TCTGGGTTTGGACATGAAATGTAAGAAAAGATTTTTCTACTGGAGCAGCTCTCTTAAGGAGAA  
ACAAATCTATTTAGAGAAAACAGCTGGCCCTGCAAAATGTTTACAGAAATGAAATCTCTCCTAC  
TTATATAAGAAATCTCACATGAGATAGAATTGTGATTTCAATAACACTTGAAAAGTGCT  
GGAGTAAACAAAATATCTCAGTTGGACCATCCTTAACCTGATTGAACTGTCTAGGAACCTTAC  
AGATTGTTCTGCAGTTCTCTCTCTTTCTCCTCAGGTAGGACAGCTCTAGCATTTCCTTAATC  
AGGAATATTGTGGTAAGCTGGGAGTATCACTCTGGAAGAAAGTAACATCTCCAGATGAGAAT  
TTGAAACAAGAAACAGAGTGTGTGAAAAGGACACCTTCACTGAAGCAAGTCGGAAGTACAA  
TGAAAAATAAATATTTTGGTATTTATTTATGAAATATTTGAACATTTTTTCAATAATTCCTT  
TTTACTTCTAGGAAGTCTCAAAGACCATCTTAAATATTATATGTTTGGACAATTAGCAAC  
AAGTCAGATAGTTAGAATCGAAGTTTTTCAAATCCATGCTTAGCTAAGCTTTTCACTCTGT  
CACTTGGCTTCGATTTTTATATTTTCTATTATATGAAATGTATCTTTTGGTGTGTTGATTT  
TTCTTTCTTTCTTTGTAAATAGTTCTGAGTTCTGTCAAATGCCGTGAAAGTATTTGCTATAA  
TAAAGAAAATCTTTGTGACTTTAAAAA

0966034-05501

# Case Study

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76400
```

><subunit 1 of 1, 487 aa, 1 stop

><MW: 55051, pI: 8.14, NX(S/T): 2

MLRAPGCLLRTSVAPAAALAAALLSSLARCSLLEPRDPVASSLSPYFGTKTRYEDVNPVLLS  
GPEAPWRDPELEGCTCTPVQLVALIRHGTRYPTVKQIRKLRLHGLLQARGSRDGGASSTGS  
RDLGAALADWPLWYADWMDGQLVEKGRQDMRQLALRLASLFPALFSRENYGRLRLITSSKHR  
CMDSSAAFQGLWQHYHPGLPPDPVADMEFGPPTVNDKLMRFFDHCEKFLTEVEKNATALYH  
VEAFKGTGPEMQNILKKVAATLQVPVNDLNADLIQVAFFTCSPDLAIKGVKSPWCDVFDIDDA  
KVLEYLNDLKQYWKRGYGYTINSRSSCTLFQDIFQHLDKAVEQKQRSQPISSPVLQFGHAE  
TLLPLLSLMGYFKDKEPLTAYNYKKQMHRKFRSLGIVPYASNLI FVLYHCENAKTPKEQFRV  
OMLLNEKVLPLAYSOETVSFYEDLKNHYKDILQSCQTSEECELARANSTSDDEL

Important features:

Signal sequence

amino acids 1-30

N-glycosylation sites.

amino acids 242-246, 481-485

N-myristoylation sites.

amino acids 107-113, 113-119, 117-123, 118-124, 128-134

Endoplasmic reticulum targeting sequence.

amino acids 484-489

# FIGURE 13

GGGACTACAAGCCGCGCGCGCTGCCCTGAGCAACCCCTCGACATGCGGCTGAGGCGGCCACCCGCGAC  
TCCGCGCTCTGCGCTGGCTGCGTACTTCTTCTGCTGCTGCTTTTCAGGGGCTGCTGATAGGGGCTGTAATC  
TCAAATCCAGCAATGAAACCCAGTGGTACAGGAATTTGAAAGTGTGGAACCTGTCTGCATCATTCAGGATTCGC  
AGCAAAGTGACCCAGGATTCGAGTGGAAAGAAATTCAGATGAACAAACCATATGTGTTTTTGACAAACAAA  
TTCAGGGAGACTTGGCGGGTCTGCAGAAATCTGGGGAAGACATCCCTGAAGATTCGGAATGTGACACGGAGAG  
ACTAGCCCTTTATCTGCTGTAGGTCTGTCTGCGAAATGACCGCAAGGAAATTTGATGAGATTGTGATTCGAGTTAA  
CTGTGCAAGTGAAGCCAGTGAACCCCTGTCTGTAGAGTGCCGAAAGCTGTACAGTAGGCAAGATGCAACACACT  
ACTGCCAGGAGAGTGTAGGGCCACCCCGGCCCTCACTACAGCTGGTATCGCAATGATGTACCACTGCCCAAGGATT  
CCAGAGCCAAATCCCAAGATTTCGCAATTTCTTTTCCACTTAAACTCTGAAACAGGCACCTTTGGTGTTCACCTGTG  
TTCACAGGAGCACTCTGGGCAGTACTACTGCATTGCTTCCAATGACGCAAGCTCAGCCAGGTGTGAGGAGCAGG  
AGATGGAAGTCTATGACCTGAACATTTGGCGGAATATTGGGGGGGTTCTGGTTGTCCTTGTCTACTGGCCCTGA  
TCACGTTGGGCATCTGCTGTGCATACAGACGTGGCTACTTTCATCAACAATAAACAGGATGGAGAAAGTTACAAGA  
ACCCAGGAAACAGATGGAGTTAACTACATCCGCACTGACGAGGAGGGCGCACTTCAGACACAAGTCATCGTTTTG  
TGATCTGAGACCCCGGGTGTGGCTGAGAGCGCACAGCGCAGCTGCACATACCTCTGCTAGAAAACCTCTGTCTCAA  
GGCAGCGAGAGCTGATGCACTCGGACAGAGCTAGACACTCACTCAGAAGCTTTTCTGTTTTGGCCAAAGTTGACCA  
CTACTCTTCTTACTTAAACAAGCCACATGAATAGAAAGAAATTTTCTCAAGATGGACCCGGTAAATATAACCAAA  
GGAAGCGAAACTGGGTGGCTTCACTGAGTTGGTTTTCTTAATCTGTGTTTGGCGCTGATTCGCGCATGAGTATTAGG  
GTGATCTTAAAGAGTTTGTCTCAGCTAAACGCCGCTGTGCGGCCCTGTGGAAGCCAGCACTGTTCAACCACTGGTCTGT  
CAGCAGCCACAGCAGCACTGTGAGATGGCGAGGTGGCTGGACAGCACCAGCGCATCCGCGGGGGAACCA  
GAAAAGGCTTCTTACACAGCAGCTTACTTTCATCGGCCACAGACACCACCGCAGTTTCTTCTTAAAGGCTCTGC  
TGATCGGTGTTGAGAGTTCATTTGGTGGAGAAGCTTTTGGATCAGCATTTTGTAAAAACAACCAAAATCAGGAAG  
GTAATTTGGTTGCTGGAAGAGGGATCTTGCTGAGGAACCTGCTTTGTCACAGGGGTGTCAGGATTATAAGGAAA  
ACCTTCGCTTTAGGCTAAGTCTGAAATGGTACTGAAATATGCTTTTCTATGGGTCTTGTTTATTTATAAAATTT  
TACATCTAAATTTTGTCTAAGGATGTATTTTGATTATTGAAAAGAAAAATTTCTATTAAACTGTAATATATTGT  
CATACATGTGTAATAACCTATTTTTTTTAAAAAGTTCAACTTAAGGTAGAAGTTCCAAGCTACTAGTGTTAAAT  
TGGAAAAATACAATAATTAAGAGTATTTTACCACAAGGAATCTCTCATGGAAGTTTACTGTGATGTTCCTTTTCT  
CACACAAGTTTTAGCTTTTTTACAAGGGAACCTATACTGTCTACACATCAGACCATGTTGCTTAGGAAACCTT  
TAAAAATCCAGTTAAGCAATGTTGAAATCAGTTTTCATCTCTTCAAAAGAAACCTCTCAGGTTAGCTTTGAAT  
GCCTCTTCTGAGATGACTAGGACAGCTGTGATCCAGAGGCCACCCAGAAGCCCTCAGATGTACATACAGAGATG  
CCAGTCAGCTCCTGGGTTGCGCCAGGCGCCCCGCTCAGCTCAGCTGTGCGCTGTCTGCCAGAGGCCCC  
GCCATCCTTGGGCCCTGSGCAGTSGCTGTGCCAGTGAGCTTTACTACGTGGCCCTGCTCTTACAGCAGCAGC  
TCTCAGGTGGGCACTGACAGGACACTGGTGTCTTCCATGTAGCGTCCCAGCTTTGGGCTCCTGTAAACAGACCTCT  
TTTTGGTTATGATGGCTCACAAAATAGGGCCCCCAATGCTATTTTTTTTTTTTAAAGTTGTTTAAATATTTGTT  
AAGATTGCTCAAGGCCAAAGGCAATTGCGAAATCAAGTCTGTCAAGTACAATAAATTTTTAAAAAGAAATGGAT  
CCCATTGTCTCTTTGCGCACAGAGAAAGCACCAGACGCCACAGGCTCTGTGCGATTTCAAAAACAAACCATGAT  
GGATGGCGGCCAGTCCAGCCTTTTAAAGAACGTCAAGTGGAGCAGCCAGGTGAAAGGCTGGCGGGGAGGAAAG  
TGAACGCCCTGAATCAAAAGCAGTTTTCTAAATTTTGAATTTTAAATTTTTCATCCGCCGGAGACACTGCTCCCAT  
TGTGGGGGACATTGACACATCACTCAGAAGCCTGTGTTCTTCAAGAGCAGGTGTTCTCAGCCTCATGTCAGCCT  
GCGGTGCTGGACTCAGACTGAAGTGTCTGTAAGCAAGGAGGTGCTGAGAGGAGCACTCACTGTGTGCGCTGGA  
GATGTGCTCTCACTACTACCTTGTCTTTCAGCTTCCAGTGTCTTGGGTTTTTTTAACTTTGACAGCTTTTTTT  
AATGTACATAGAGACTGTGTTGACTTTTTTAAAGTATGATGTAACACTTTGCGCAGGCGCGCTGCGAGAGGCA  
GGAAATGCTTCAGCAGTGGCTCAGTGTCTCCTGGTGTCTGTCATGCACTCTGATGCTTAGCATTCGCAAGTTC  
CCTCCATCATTGCCACCTTGGTAGAGAGGATGGCTCCCCACCTCAGCCTTGTGGGATTCACGCTCCAGCCTCT  
TCTTGGTGTGCTAGTGTAGAGGTAGGCTTATTTGCCCTCTTCTTATACCTTAAACCTCTACACTAGTGCCA  
TGGGAACCAAGTCTGAAAGAGTAGAGAGAAGTGAAGTAGAGTCTGGGAAGTAGCTGCTTAACTAGCAGACTAGA  
CGGAAAGGAATATCTCGTGTATTTTAAAGATGAACTCAAGACTCGAGCGGATACGAGCTCGGCTGTGATTTCT  
GCCTTTGGATGGATGTTGCTGTACACAGATGCTACAGACTGTGACTAACACACCGTAAATTTGGCATTGTTTAAAC  
CTCATTTATAAAGGCTTCAAAAAACCCA

0966034-052501

[illegible]

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77624
```

```
><subunit 1 of 1, 310 aa, 1 stop
```

><MW: 35020, pI: 7.90, NX(S/T): 3

MALRRPPRLRLCARLPDFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTS  
PRIEWKKIQDEQTTYVFDNKIQGLAGRAEILGKTSLKIWNVTRRDSALYRCEVVARNRK  
EIDEIVIELTVQVKPVPVCRVPAKVPVGKMATLHCQSEGHPRPHYSWYRNDVPLPTDSRA  
NPRFRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVDNLNIGGIIGG  
VLVVLAVLALITLGCAYRGRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSSFI

Important features of the protein:

Signal peptide:

amino acids 1-30

## Transmembrane domain:

amino acids 243-263

N-glycosylation sites.

amino acids 104-107, 192-195

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 107-110

Casein kinase II phosphorylation site.

amino acids 106-109, 296-299

Tyrosine kinase phosphorylation site.

amino acids 69-77

N-myristoylation sites.

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267

# FIGURE 15

CAGGACCAGGTCCTTCTACGCTGGAGCAGCGGGAGACAGCCACCATGACACATCTCTGTTGTCATGCCATGGTG  
ATCCTGCTGACGCTGGGGCCGCTCTCGAGCCGACGACAGCGAGTTCAGGGCGCTGCTGGACATCTGGTTTCGGAG  
GAGAGGACCATGCCACCGCTCTCTGGTGGACACATCGGAGGAGGCGCTGCTGCTTCTGACTGGCTGAAGCTG  
CGCATGATCCGTTCTAGGCTGCTCCGCTGGTGGACGCGCCCTGACGACCTGGAGCCGACGACGCTGCTGCTG  
TTCTGTCAGTCTGTTTGGCATCCCCGTGTCCAGCATGAGCAAACTCTCTCAGTTTCCTGACACAGGACAGTGGCCAC  
GACCCCGACATCTGGAGCAGAACTCATGGACAAGAATTACATGGGCCACCTGGTGGAGGTCCAGCATGAGCGC  
GGCGCTCTCGGAGGCGACACTTTCCACTCTCTGCTCACAGGCTCTCCCTCGCCGCCGCGGAGACAGCACAGGGCA  
CCCAAACCAAAGAGCAGCCAGAGCAGCCATAGGCGAGGGCGGATTTCGGGTGGGGACCCAGCTCCGGGTGCTG  
GGCCCTGAGGACAGACTGGCTGGCATGTCTCTCCAGATTTTCCGCTCAGCCGAGACCCCTCGGTGGCAGAGCTCC  
AGTCCCCGCCCTGCGCCTCGCCCTGACAGGCGCTGGGCGAGGAGCTGGCCGCGCTCGTCCAGGGCAGCCCC  
GAGGTGCGGGGCATACGGTGGTGTCTGTCAGGCGCTCGCCACCCCTGCTCAGCTCCCCACAGCGCGGTGCTCTG  
GTGATGTCCATGCAACCTAGCCACTTCTGGCCTGCGCCGTGCTGCGCCAGCTCTGCCAGTACACGCGCTGTGTG  
CCACAGGACACCGGCTTCTCTGCTCTTCTGAAAGGTGCTCTGTCAGATGCTGCAAGTGCTGGACAGCCCTGAC  
GTGCGAGGGCGGCCCTGCGGGCACAGCTCAGGATGCTTGCAGCCAGGCGCTCAGCCGGCGAGGCTCAGTGTAT  
GTGCGAGGGGGCTCTCTGCGCTGGCGAGGCGCTGACCTTCCGTGAGGACCTGGAGGTGGTCACTCCACCGTCT  
CGTGGCTCATCGCCACCTGAGGTCTGGGGAGCAGTGACGCTGGAGCGGAGCTGATCAGCAAGTCTCTCCAG  
GGCTGATCGAGGTGAGGTTCGCCACCTGGAGGAGCTGCTGACTGATTCTTCTGCGCACTGGCGGATGCTGTC  
TCCCGTTTTCAGCCTGTAAGCCCGTTGTGGTGGTGAAGTCTCTGCTGCTGAGGAGGAGGAGCCCTGGCTGGG  
GGGAGGCTGGGTGCGGACGCTGGCGCACAGCTGAGGCGGTGGGCGCTCTGTCAGGCTCTCAGTGGCTCAGTGG  
CTGGAATGCTGGAACCCGAGGTGGTACAGAGCTGCCCCGACCTGACGCTCAGGCTGCTCTTCTCCCGAGGAAG  
GGCAAAAGTCAAGGCCAGGTGCCCTCGTTCTCGTCCCTACTCTGACCTCTTCAAGCATCAGTCCAGCTGGCCCC  
ACACTGACCAAGTGATCCGAGTCTCTGCTGGCAAGAGCCGGAAACAGAGGTTCGACCCCTCTGCTCTCTGGAC  
TTCTCTGGCGCTGCATCCATGTTCTCTGCACTGCGAGGGCGGGACAGCGCACCCCGCAGAGCGCGGGAG  
GAGCTGGTCTGCGGGTCCAGGGCCCGGAGCTCATCAGCTGGTGGAGCTGATCTGGCCCGGGAGGACGCGG  
AGCCAGGACCGGGACACAGCGCGCTGAGCCTCATCAGGCCCGGCTGCCCCGTGCTGCTCAGTGTGCTGTGGG  
GACGATGAGAGTGTGAGAAAGTGACGAGCACCTGTGAGGCTGCATCAGAGCTGGGAGACAGCGCTGCTGGGA  
AGGCGCTGCGAGACCTTCTCTGCAAGCACTACCTACAGCGCGCGGAGCTGCGGGTGCCCGTGCCTGAGGTCTTA  
CTGCACAGCAAGGGCTGCCAGCAGCAGCGTCTGCAAGCTGGACGGAATCATCACCGCTTCATCAGCTCCTT  
GCGGACACAGCGACTCCCGGGCTTGGAGAACGAGGGCGGATGCCAGCATGGCGCTGCGGGAAGCTGGCGGTG  
GCGCACCCGCTGCTGCTGCTCAGGCACCTGCCATGATCGCGCGCTCTGTCAGCGCGCACCCACCTCAACTTC  
CAGGATTTCCGGCAGCAGAACCACTGAGCTGCTTCTGCACTGCTGGGCTGTGGAGCTGCTGACGCGCGCAC  
GTGTTCCGAGCAGCACACGCGGGCGCTGTGGAGTGCCTTCTGCTCTCATCCGCTGCTGCTGAATTACAGG  
AAGTCTCTCCGCCATCTGGCTGCCTCATCAACAAGTTTGTGCAAGTTCATCCATAGTACATTACCTACAATGCC  
CCAGAGCCATCTCTTCTCTGCAAGACACGCGGACCGCTCCAGACCTGCTCTCGACAAACAGTGACCTGGTG  
ATGCTGAAATCCCTCTCTGAGGGCTCAGCCTGCCAGCAGGAGCAGCAGGACCGACCGAGGCTGGAGCAAGAG  
GGCAGGAGGAGAGCTCAGCGGCTCTTGGCCCTGGTCAGCGTCTCTGTTTCACTCCCTGACCGCGGCCGAG  
ATGGCCCCCTCATGAAACGCTTTCCTGGGGCCAAACGGTGGAGGATCTGCTGGAGGTTCTGAGTGACATAGAC  
GAGATGTCCCGCGGAGACCCGAGATCTGAGCTTCTTCTCGACCAACCTGATGAGCGCTGATGAGCTGGGCCGAG  
GAGTGTGTCGCAACCTCGGCTCTCAGCTGGCCCTGCGCTCGAGCAACAGCCCCAGCATTGCGAGCGCTTCT  
CTGCCACCTGCTCATGACTGCTGGGACGACGAGACTTTGAGTGGTGACAGCGGCCCTCGCGAAGCTGCTGAG  
TAGCTCTCTGTGTCGAAGAGCAGCGGCTGTGCTGCTCCACCGGCGCTCTGAGTGGGCACTGTACGCGCAGATG  
GACCCAGCGCGCAGATCTCCGAGGCGCTGAGGATCTGCATATGAGGCGCTGATGTGAGCTGTGGGACCGCA  
CCCCCTCCAAGCCCCGCGCTGCTGCTCCCGGGATCTCGAGGCCAGGCGCCAGGAGCAATACTCCGAGCCCTGGGTGG  
TCTGTCCGAGGAGTGAGGGCGAGCCCTGAGGCCAGGCGCCAGGAGCAATACTCCGAGCCCTGGGTGG  
CTCCGGGCGCGCGCTGATCATAGGGCGCTCCAGCAAGCCCTCATTCACCTTCTGGGCCACAGCCCTGCCCGG  
AGGCGGGATCCCCCGGCGCTGCTGGCTGGTTTGAATGAACACGACTGAATGTCTCA

0966034.052501

## FIGURE 16

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77631

><subunit 1 of 1, 1029 aa, 1 stop

><MW: 114213, pI: 6.42, NX(S/T): 0

MHILVVHAMVILLTLGPPRADDSFQALLDIWFPBEKPLPTAFLVDTSEALLLPDWLKLRM  
IRSEVLRLVDAALQDLEFPQQLLFVQSFQIPVSSMSKLLQFLDQAVAHDPQTLEQNIIMDKNY  
MAHLVEVQHERGASGGQTFHSLLTASLPPRRDS TEAPKPKSSPEQPIGQGRIRVGTQLRVLG  
PEDDLAGMFLQIFPLSPDRWQSSSPRPVALALQQALGQELARVVQGSPEVPGITVRVLQAL  
ATLLSSPHGGALVMSMRSHFLACPLLRLQCQYQRCVPQDTGFSSFLFKVLLQMLQWLDS PG  
VEGGPLRAQLRMLASQASAGRRLSDV RGGLRLAEALAFRQDLEVVSSTVRAVIATLRSGEQ  
CSVEPDLISKVLQGLIEVRSPHLELLTAFFSATADAASPPACKPVVVVSSLLQEEEP LA  
GGKPGADGGSLEAVRLGPSSGLLDVWLEMLDPEVVSSCPDLQLRLLFSSRRKGGQAQVPSFR  
PYLLTLFTHQSSWPTLHQCIQVLLGKSREQRFDPSASLDLFWACIHVPRIWQGRDQRT PQKR  
REELVLRVQGP ELISLVELILAEAE TRSQDGD TAACSLIQARLP LLLSCCGDDSVRKVTE  
HLSGCIQQWGD SVLGRRCRDLLQLYLQRP ELRVVPVEVLLHSEGAASSSVCKLDGLIHRFI  
TLLADTSDSRALENRGADASMACRKLAVAHPLLLLRHLP MIAALLHGRTHLNFQEF RQQNHL  
SCFLHVLGLLELLQPHVFRSEHQGALWDCLLSFIRLLLN YRKSSRHAAFINKFVQFIHKYI  
TYNAPAAISFLQKHADPLHDLSDFNSDLVMLKSLLAGLSLPSRDDRTDRGLDEEGEEBESSAG  
SLPLVSVSLFTPTLTAAEMAPYMKRLSRGQTVEDLLEVLSDIDEMSRRRRPEILSFSTNLQRL  
MSSAECCRNLAFSLALRSMQNSPSIAAFLPTFM YCLSGSQDFEVVQTALRNLPEYALLCQE  
HAAVLLHRAFLVGMYGQMDPSAQISEALRILHMEAVM

### Important features:

#### Signal peptide:

amino acids 1-16

#### cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 154-158, 331-335, 616-620, 785-789, 891-895

#### N-myristoylation sites.

amino acids 91-97, 136-142, 224-230, 435-441, 439-445, 443-449,  
665-671, 698-704

#### Amidation sites.

amino acids 329-333, 634-638

0306034-052501



# FIGURE 17

CCGGGCCATG CAGCCTCGGCCCGCGGGCGCCCGCCGCGCACCCGAGGAGATGAGGCTCCGC  
AATGGCACCTTCTTGACGTGCTGCTCTTCTGCTGTGCGCCTTCCCTCTCGTGTCTCTGGTA  
CGCGGCACCTCAGCGGCCAGAAAGGCGACGTTGTGGACGTTTACCAGCGGGAGTTCCTGGCGC  
TGCGCGATCGGTTGCACGAGCTGAGCAGGAGAGCCTCAAGCGCTCCAAGGAGCTCAACTG  
GTGCTGGACGAGATCAAGAGGGCCGTGTGAGAAAGGCGGCTGCGAGACGGAGACGGCAA  
TCGACCTGGGGCCGCTTAACAGAGGACCCCGATTGAAGCCGTGGAAACGGCTCACACCGGC  
ACGTGCTGCACCTGCCACCGTCTTCCATCACCTGCCACACCTGCTGGCCAAAGGAGAGCAGT  
CTGCAGCCCGCGGTGCGCGTGGGCCAGGGCCGACCGGAGTGTGCGTGGTGATGGGCATCCC  
GAGCGTGCGGCGGAGGTGCACTCGTACCTGACTGACACTCTGCACTCGCTCATCTCCGAGC  
TGAGCCCGCAGGAGAAGGAGGACTCGGTGCTGCTGGTGCTGATCGCCGAGACTGACTCACAG  
TACACTTCGGCAGTGACAGAGAACATCAAGGCCTTGTTCCCCACGGAGATCCATTCTGGGCT  
CCTGGAGGTGCTCATCTACCCCTCCCCCACTTCTACCTGACTTCTCCCGCTCCGAGAGTCTT  
TTGGGACCCCAAGGAGAGAGTCAAGGTGGAGGACCAACAGAACCTCGATTACTGCTTCTCTC  
ATGATGTACGCGAGTCAAAGGCATCTACTAGTGCAGCTGGAGGATGACATCGTGGCCAA  
GCCCAACTACCTGAGCACCATGAAGAACTTTGCACTGCAGCAGCCTTCAGAGGACTGGATGA  
TCCTGGAGTTCTCCAGCTGGGCTTCATTGGTAAGATGTTCAAGTTCGCTGGACCTGAGCCTG  
ATTGTAGAGTTCACTTCTCATGTTCTACCGGACAAAGCCCATCGACTGGCTCCTGGACCATAT  
TCTGTGGGTGAAAGTCTGCAACCCCGAGAAGGATGCGAAGCAGTGTGACCGGCAGAAAGCCA  
ACCTGCGGATCCGCTTCAAACCGTCCCTCTTCCAGCACGTGGGCACTACTCTCGCTGGCT  
GGCAAGATCCAGAACTGAAGGACAAAGACTTTGGAAGCAGGCGCTGCGGAAGGAGCATGT  
GAACCCGCCAGCAGAGGTGAGCAGGAGCTGAAGACATACCAGCACTTCACCTGGAGAAAG  
CCTACCTGCGCAGGAGCTTCTTCTGGGCTTACCCCTGCCGCGGGGGACTTCATCCGCTTC  
CGCTTCTTCCAACTCTAAGACTGGAGCGGTTCTTCTCCGAGTGGGAACATCGAGCACCC  
GGAGGACAAAGCTCTTCAACAGTCTGTGGAGGTGCTGCCCTTCGACAACCCCTCAGTCAGACA  
AGGAGGCCCTGCGAGGAGGCCGACCCGCCACCTCCGGTACCCCTCGGAGCCCCGACGGCTAC  
CTCCAGATCGGCTCCTTCTACAAGGAGTGGCAGAGGAGAGGTGGACCCAGCCTTCGGGCC  
TCTGGAAGCACTGCGCTCTCGATCCAGACGGACTCCCTGTGTGGGTGATTCTGAGCGAGA  
TCTTCTGAAAAAGGCGCACTAAGCTGCGGGCTTCTGAGGTACCTGTGGCCAGCCCTGAA  
GCCACATTTCTGGGGTGTGCTACTGCCGTCCCCGAGGGCCAGATACGGCCCGGCCCAA  
AGGGTTCTGCCTGGCGTCGGGCTTGGGCCGGCTGGGGTCCGCCGTGGCCCGGAGGCCCTA  
GGAGCTGGTGCTGCCCCGCCGCCCGGGCCGCGGAGGAGGCAGGCGGCCCCACACTGTGTC  
TGAGGCCCGGAACCGTTCGCACCCGGCTGCCCCAGTCAGGCCGCTTTTAGAAGAGCTTTTAC  
TTGGGCGCCCGCGTCTCTGGCGGAACTGGAATGCATATACTACTTTATGTGCTGTGTT  
TTTTATTCTTGATACATTTGATTTTTTACGTAAAGTCCACATATACTTCTATAAGAGCGTG  
ACTTGTATAAGGGTTAATGAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAA

0966034.052501

## FIGURE 18

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA82307
><subunit 1 of 1, 548 aa, 1 stop
><MW: 63198, pI: 8.10, NX(S/T): 4
MRLRNGTFLTLLLFCLCAFLSLSWYAALSGQKGDVVDVYQREFLALRDLHAAEQESLKRSK
ELNLVLDEIKRAVSEKALRDGDGNRTWGRLTEDPRLKPNWNGSHRHVLHLPVTFHHLPHLLA
KESSLQPAVRVGGQRTGVSVVMGIPSVRREVHSYLTDTLHSLISELSPQEKEDSVIVVLIAE
TDSQYTSAVTENIKALFPTEIHSGLLEVISPSPHFYPDFSRLRESFGDPKERVVRWRTKQNLN
YCFLMMYAQSKGIYYVQLEDDIVAKPNYLSTMKNFALQQPSDWMILEFSQLGFIGKMFKSL
DLSLIVEFILMFYRDKPIDWLLDHILWVKVCNPEKDAKHCDRQKANLRIRFKPSLFQHVGTN
SSLAGKIQKLKDKDFGKQALRKEHVNPPAEVSTSLKTYQHFTLEKAYLREDDFFWAFTPAAGD
FIRFRFFQPLRLERFFFRSGNIEHPEDKLFNTSVEVLPPDNPPQSDKEALQEGRATLRYPRS
PDGYLQIGSFYKGVAEGEVDPAFGPLEALRLSIQTDSPVWVILSEIFLKKAD
```

### Important features:

#### Signal sequence

amino acids 1-23

#### N-glycosylation sites.

amino acids 5-9, 87-91, 103-107, 465-469

#### N-myristoylation sites.

amino acids 6-12, 136-142, 370-376, 509-515

0906034.052501